40

METHODS OF RESPONSE SURFACE ANALYSIS

by

JOYCE LYNN TAYLOR

B. S., Oklahoma State University, 1973

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1975

Approved by:

George A. Milliken
Major Professor

TABLE OF CONTENTS

ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

INTRODUCTION

In many fields of research an experimenter has the task of finding the best set of operating conditions for a system which could have many different factors each at several levels. The system he is optimizing may be a chemical experiment with many chemicals at different levels of concentration or a machine with different settings affecting production rate or a storage container with different factors influencing its performance. In any field, once the researcher has gathered this data, one type of analysis he can perform is response surface analysis.

The purpose of this paper is to describe the analysis of a response function and illustrate the computational techniques needed with examples. The analysis consists of three parts: (1) the regression analysis which fits the response surface to the data and locates the stationary point where the best set of operating conditions may occur, (2) the canonical analysis which yields information about the nature of the surface, and (3) the determination of a confidence region for the location of the stationary point.

The response function which will be analyzed is a second-order model with k independent variables of the form

$$\eta = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{\substack{j \\ j<m}} \sum_{m} \beta_{jm} x_j x_m + \sum_{j=1} \beta_{jj} x_j^2$$

since a second-order model is usually assumed to be sufficient to describe a response surface in the region of a stationary point. The examples which follow the discussion of the analysis use a second-order model with two independent variables. Output from a computer program which performs the three parts of the response surface analysis and plots the confidence region for the stationary point using CALCOMP plotting routines is also included.

COMPUTATIONAL TECHNIQUES

REGRESSION ANALYSIS: To estimate the β-coefficients of the response function $\eta = \beta_0 + \sum_j \beta_j x_j + \sum_{\substack{j \\ j<m}} \sum_m \beta_{jm} x_j x_m + \sum_j \beta_{jj} x_j^2$ with k independent variables, the relationship is restated in matrix notation as $\underline{y} = \underline{X\beta}$, which can be expanded to show the vector elements as

$$\begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{matrix} = \begin{matrix} 1 & x_{11} & \cdots & x_{k1} & x_{11}^2 & \cdots & x_{k1}^2 & x_{11}x_{21} & \cdots & x_{k-1,1}x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} & x_{12}^2 & \cdots & x_{k2}^2 & x_{12}x_{22} & \cdots & x_{k-1,2}x_{k2} \\ \\ 1 & x_{1n} & \cdots & x_{kn} & x_{1n}^2 & \cdots & x_{kn}^2 & x_{1n}x_{2n} & \cdots & x_{k-1,n}x_{kn} \end{matrix} \quad \begin{matrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \\ \beta_{11} \\ \vdots \\ \beta_{kk} \\ \beta_{12} \\ \vdots \\ \beta_{k-1,k} \end{matrix}$$

where the n observed responses of the experiment are the $\underline{y}$-vector. Then $\underline{b} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$ is the vector of estimates of the β's, {3}

Once the β's have been estimated, the fitted response surface can be described by the relation

$$\hat{y} = b_0 + \sum_j b_j x_j + \sum_{\substack{j \\ j<m}} \sum_m b_{jm} x_j x_m + \sum_j b_{jj} x_j^2$$

or in matrix notation as

$$\hat{y} = b_0 + \underline{x}'\underline{b}_1 + \underline{x}'\underline{Bx}$$

where now $\underline{b}$ has been separated into the first order terms of $\underline{b}_1$ and the mixed quadratic (b_{ij} where $i \neq j$) and pure quadratic (b_{ii}) coefficients of the matrix $\underline{B}$ so that

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \quad \underline{b}_1 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}, \text{ and } \underline{B} = \begin{bmatrix} b_{11} & b_{12}/2 & \cdots & b_{1k}/2 \\ & b_{22} & \cdots & b_{2k}/2 \\ & & & \vdots \\ & & & b_{k-1,k}/2 \\ & & & b_{kk} \end{bmatrix}.$$

The following table shows the degrees of freedom (dof) associated with the sources of variation in this analysis.

Source	dof
Mean (β_0)	1
Regression	
Linear	k
Quadratic	$\frac{k(k+1)}{2}$
Residual	$n-(k+\frac{k(k+1)}{2}+1) = \phi$

This table will be useful when the boundary of the confidence region about the stationary point is being determined.

If multiple observations are taken at the points $\underline{x}_i = (x_1,\ldots,x_k)_i$ a measure of the lack of fit of the fitted surface to the second-order response model can be made. The degrees of freedom for residuals can be divided into degrees of freedom for pure error and degrees of freedom for lack of fit.

Residual		
Lack of Fit	(by subtraction: ϕ - P.E. dof)	
Pure Error	$\sum_{i=1}^{t}(n_i - 1)$	where t is the no. of points replicated

A test of the lack of fit could be made by comparing the test statistic Lack of Fit Mean Square/Pure Error Mean Square to an F-value with degrees of freedom for the numerator equal to Lack of Fit dof and degrees of freedom for the denominator equal to Pure Error dof.

Next the stationary point of this surface can be defined as the solution of $\frac{\partial \underline{y}}{\partial \underline{x}} = \frac{\partial}{\underline{x}}(\underline{x}'\underline{b}_1 + x'Bx) \overset{set}{=} \underline{0}$

$$\underline{b}_1 + 2\underline{Bx} = \underline{0}$$

$$\underline{x}_0 = -\underline{B}^{-1}\underline{b}_1/2 \ .$$

Where $\underline{x}_0$ is the stationary point. This stationary point may be an optimum (maximum or minimum) or it may be a saddle point of the response surface. So the stationary point has been found as a function of the estimated regression coefficients,{4}.

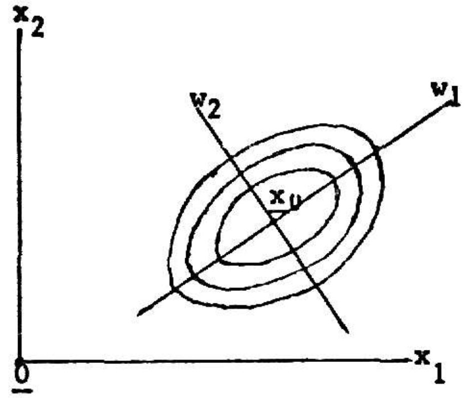
CANONICAL ANALYSIS: If the response surface can be plotted, one does not have too much trouble visualizing the nature of the surface. However, a second-order model in k independent variables requires a (k+1)-dimensional plot which would be difficult to evaluate.

The canonical analysis depends upon the characteristic roots and vectors of the matrix $\underline{B}$. The characteristic roots of $\underline{B}$, denoted by $\lambda_1, \lambda_2, \ldots, \lambda_k$, can be found as the solution to $|\underline{B} - \lambda_i \underline{I}_k| = 0$ and the characteristic vector $\underline{m}_i$ associated with the characteristic root λ_i is the solution vector to $(\underline{B} - \lambda_i \underline{I}_k)\underline{m}_i = \underline{0}$.

The nature of the stationary point can be determined by examining the characteristic roots. If all the roots are positive the stationary point is a minimum, and if all the roots are negative then $\underline{x}_0$ is a maximum. In the case where the λ's differ in sign, the stationary point is a saddle point.

The magnitude of the λ's also adds information about the response surface. If $|\lambda_i|$ is considerably smaller than the other characteristic roots, then the contour system of the surface is elongated in the direction associated with λ_i. This implies that a change in this direction would not greatly affect response, and one might have a range of possible values which give responses very near the response at $\underline{x}_0$.

Geometrically the canonical analysis is a translation of the origin accompanied by an axis rotation. The figure on the next page illustrates the situation for two variables, x_1 and x_2. The origin is translated from $\underline{0}$ to $\underline{x}_0$ and the original x_1-x_2 axis system is rotated into a new position corresponding to the w_1 and w_2 axes in the figure.



If $|\lambda_2| >> |\lambda_1|$ then, as the figure shows, the contour system would be elongated along the w_1 axis and responses would not greatly change along this axis.

The old x-variables are related to the new w-variables by the orthogonal transformation $\underline{M}'\underline{z} = \underline{w}$ where $\underline{M}$ has as its columns the normalized characteristic vectors of $\underline{B}$ and $\underline{z} = \underline{x}-\underline{x}_0$. This shows the orientation of the contour system of the surface with respect to the x-axis system since the major and minor axes of the ellipsoidal contours are the w axes. In terms of the w-variables, the fitted response function $\hat{y} = b_0 + \underline{x}'\underline{b} + \underline{x}'\underline{Bx}$ can be written as

$$\hat{y} = b_0 + \underline{x}_0'\underline{b} + \underline{x}_0'\underline{Bx}_0 + \lambda_1 w_1{}^2 + \lambda_2 w_2{}^2 + \ldots + \lambda_k w_k{}^2$$

$$= \hat{y}_0 + \lambda_1 w_1{}^2 + \ldots + \lambda_k w_k{}^2$$, the canonical form of the response function, {4} and {5}.

A CONFIDENCE REGION FOR THE STATIONARY POINT: The confidence region for the stationary point of the response surface is the confidence region for the solution of a set of k simultaneous equations

$$\frac{\partial \underline{y}}{\partial \underline{x}} = \underline{b}_1 + 2\underline{B}\underline{x} = \underline{\delta} \quad \text{where } E(\underline{\delta}) = \underline{0} \text{ and } Var(\underline{\delta}) = E(\underline{\delta}\underline{\delta}') = \sigma^2\underline{V},$$

as given in {1}. The estimate of σ^2 is the Residual Mean Square = s^2 which is distributed $\chi^2(\phi)$. $\underline{V}$ can be calculated since $Var(\underline{b}) = \sigma^2(\underline{X}'\underline{X})^{-1}$ has already been determined.

Since $\dfrac{\underline{\delta}'\underline{V}^{-1}\underline{\delta}}{ks^2}$ is distrubuted as $F(k,\phi)$, the probability is $1-\alpha$ that the inequality $\dfrac{\underline{\delta}'\underline{V}^{-1}\underline{\delta}}{ks^2} < F_\alpha(k,\phi)$ is true. Then the boundary of the exact $(1-\alpha)100\%$ confidence region is given by

$$\underline{\delta}'\underline{V}^{-1}\underline{\delta} = ks\ F(k,\phi).$$

Another, more readily computable form for the boundary of the confidence region can be found using this result. First observe that for any matrix $\underline{Z} = \begin{bmatrix} \underline{Z}_{11} & \vdots & \underline{Z}_{12} \\ \cdots & \vdots & \cdots \\ \underline{Z}_{21} & \vdots & \underline{Z}_{22} \end{bmatrix}$; $|\underline{Z}| = |\underline{Z}_{22}| \cdot |\underline{Z}_{11} - \underline{Z}_{12}\underline{Z}_{22}^{-1}\underline{Z}_{21}|$.

Let $\underline{Z}_{11} = 0$, $\underline{Z}_{12} = \underline{\delta}'$, $\underline{Z}_{21} = \underline{\delta}$, and $\underline{Z}_{22} = \underline{V}$

then
$$\begin{vmatrix} 0 & \vdots & \underline{\delta}' \\ \cdots & \vdots & \cdots \\ \underline{\delta} & \vdots & \underline{V} \end{vmatrix} = |\underline{V}| \cdot |-\underline{\delta}'\underline{V}^{-1}\underline{\delta}| = |\underline{V}| \cdot (-\underline{\delta}\ \underline{V}^{-1}\underline{\delta}).$$

Now the original quadratic form can be written as a ratio of determinants

$$\underline{\delta}'\underline{V}^{-1}\underline{\delta} = -\frac{\begin{vmatrix} 0 & \vdots & \underline{\delta}' \\ \cdots & \vdots & \cdots \\ \underline{\delta} & \vdots & \underline{V} \end{vmatrix}}{|\underline{V}|}.$$

Now the boundary of the confidence region is defined by those values of $x_1, \ldots, x_k$ which satisfy

$$\left| \begin{array}{c:c} ks\, F(k,\phi) & \underline{\delta}' \\ \hdashline \underline{\delta} & \underline{V} \end{array} \right| = 0 \quad ,(1).$$

EXAMPLES

The examples in this section use a second order response model with two independent variables. Example 1 has nine coded observations arranged in the manner of a central composite design. The effect of several different α-levels on the confidence region is shown using a CALCOMP plotting routine which plots the confidence region for the stationary point of the response surface. Example 2 uses the same nine observations and adds six observations at the origin so that a measure of the lack of fit of the model can be made.

For both examples the true response model is taken to be

$$\eta = 78.373 + 4.533x_1 - 1.867x_2 - 3.333x_1^2 - 3.333x_2^2 + 4.00x_1x_2.$$

Data was generated by adding random normal deviates mean 0 and variance 1 to η so that $y = \eta + \varepsilon$.

The nine observations generated are as follows:

x_1	x_2	η	ε	y
1	1	78.373	-.381	77.992
1	-1	74.107	1.592	75.699
-1	1	61.307	.034	61.341
-1	-1	73.041	.573	73.614
0	1.414	69.069	.175	69.244
0	-1.414	74.349	.999	75.348
1.414	0	78.119	2.083	80.202
-1.414	0	65.299	.475	65.774
0	0	78.373	-.117	78.156

With source of variation table

Source	dof
Mean	1
Linear	2 = k
Quadratic	3
Residual	3 = ϕ
Total	9

Example 1 a: A computer program which performs the three parts of a response surface analysis was used to analyze the data just given. The parameter cards used in the program were:

```
(3F10.3)            95% C.R. (9 OBS.)            -X1 AXIS-X2 AXIS-
  9 -1 17    9.5          9.55        -5.   5.   -5.   5.
```

The value of FVAL if F(2,3) = 9.55 so a 95% confidence region is plotted.

The first page of results which follow is an acho check of the parameter cards just displayed. Next are the regression and canonical analyses. The regression coeffieients give the fitted response function as $\hat{y} = 78.156 + 4.893x_1 - 2.327x_2 - 2.705x_1^2 - 3.051x_2^2 + 3.64x_1x_2.$

The $(\underline{X}'\underline{X})^{-1}$ matrix is shown as well as the stationary point $\underline{x}_0$ = (1.0829, .26495) which is a maximum since both characteristic roots are negative. The value Turing's conditioning measure which is given is calculated as

$$C = \sqrt{\frac{\sum_{i=1}^{k} \lambda_i \sum_i (1/\lambda_i)}{k}}$$

and is a measure of how elongated the confidence region is. If $1 \leq C \leq 1.5$ then the region is acceptable. However if C>1.5 the region is severely elongated implying there are many "nearly correct" solutions to the equation $\dfrac{\partial \underline{y}}{\partial \underline{x}} = \underline{b} + 2\underline{B}\underline{x} = \underline{0}$.

On the second page of output is also shown the value ks^2F which defines the boundary of the confidence region. The third page is the plot of the 95% confidence region. There the original axes are represented by the dashed lines while the canonical axes are the solid lines. Observations within the range of the plot are shown as x's. The confidence region boundary is plotted continuously. Observe that the ellipsoidal confidence region leaves the first quadrant "wraps around" infinity and appears again in the fourth quadrant.

One should take note that this and the other confidence regions which are shown are confidence regions for the location of the stationary point $\underline{x}_0$ and not for responses at $\underline{x}_0$. The confidence region does to some extent reflect the shape of the contour system of the surface, however.

ECHO CHECK OF PARAMETER CARDS

N	IOUP	NCHAR	PLTH	FLOFT	FVAL	FPCCL	X1-AXIS PLOT LIMITS		X2-AXIS PLOT LIMITS	
9	-1	17	9.500	9.550	9.550	0.0	-5.000	5.000	-5.000	5.000

DATA INPUT FRMT	PLOT TITLE	X1AXIS LBL	X2AXIS LBL
(3F10.3)	95% C.R. (9 OBS.) -	X1 AXIS-	X2 AXIS-

95% C.R. (9 OBS.)

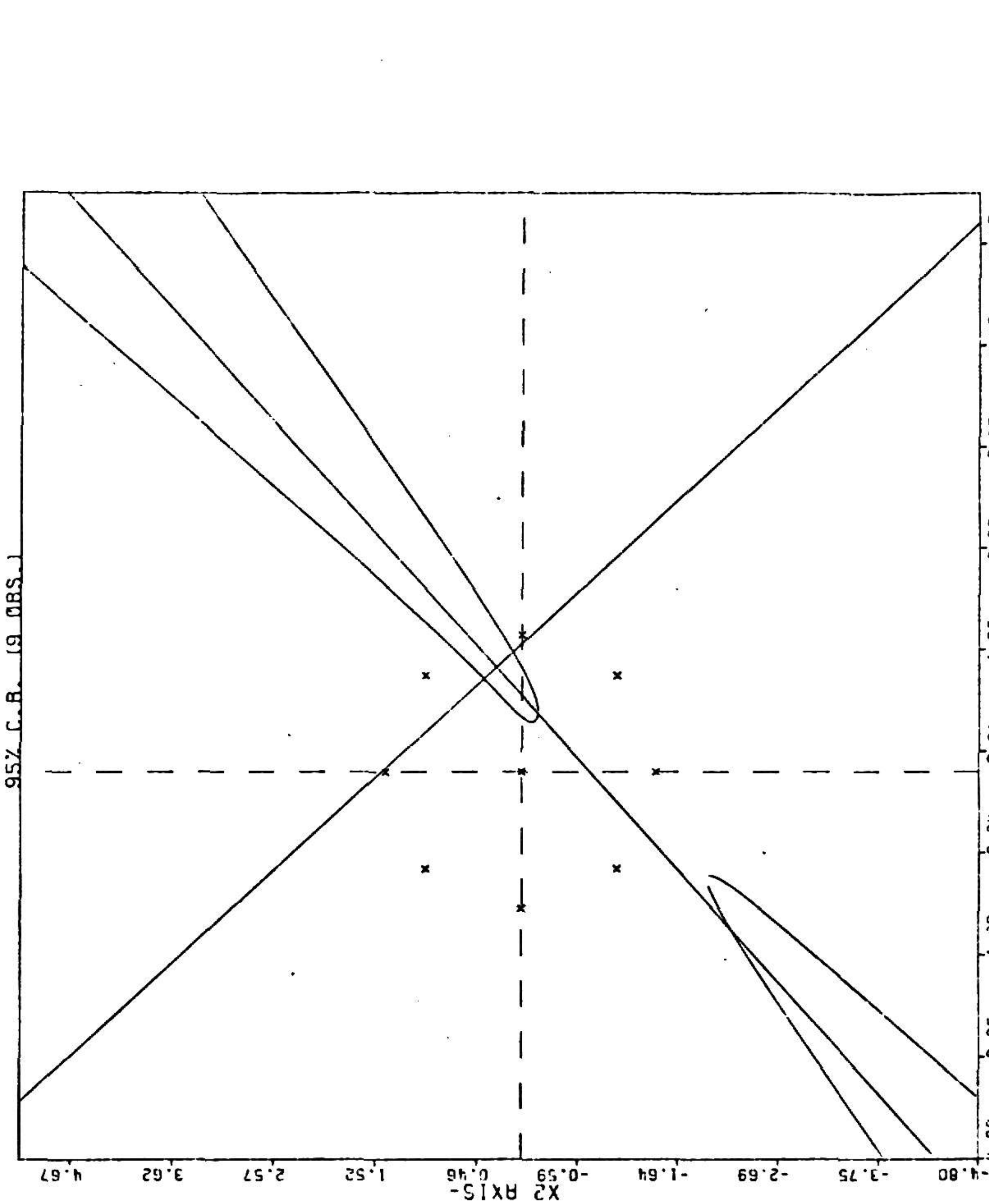
ERROR VARIANCE USED IS 0.344658 TO GIVE K*F*S**2 = 6.58374

REGRESSION COEFFICIENTS	0	1	2	11	22	12
	78.1563	4.85299	-2.32673	-2.70467	-3.05077	3.64150

INVERSE MATRIX	1	2	11	22	12
1	0.125019				
2	0.0	0.125019			
11	0.0	0.0	0.343873		
22	0.0	0.0	0.213797	0.343873	
12	0.0	0.0	0.0	0.0	0.250000

TURINGS CONDITIONING MEASURE C = 1.295246

STATIONARY POINT		EIGENVALUES	EIGENVECTORS	
X(1)=	1.0829	-1.0488	0.73980	0.67282
X(2)=	0.26455	-4.7067	-0.67282	0.73980


95% C.B. (9 OBS.)
X1 AXIS-
X2 AXIS-
-4.00
-2.95
-1.89
-0.84
0.21
1.26
2.32
3.37
4.42
5.47
-4.80
-3.75
-2.69
-1.64
-0.59
0.46
1.52
2.57
3.62
4.67

Examples 1 b and 1 c: The next two examples illustrate the confidence region at two other probability levels. The same nine observations are used. For example 1 b. where a 99% confidence region is plotted, FVAL = $F_{.01}(2,3)$ = 30.8 and the boundary is defined be ks^2F = 21.2334. For example 1 c. which is a plot of a 75% confidence region, FVAL = $F_{.25}(2,3)$ = 2.28 and ks^2F = 1.57182.

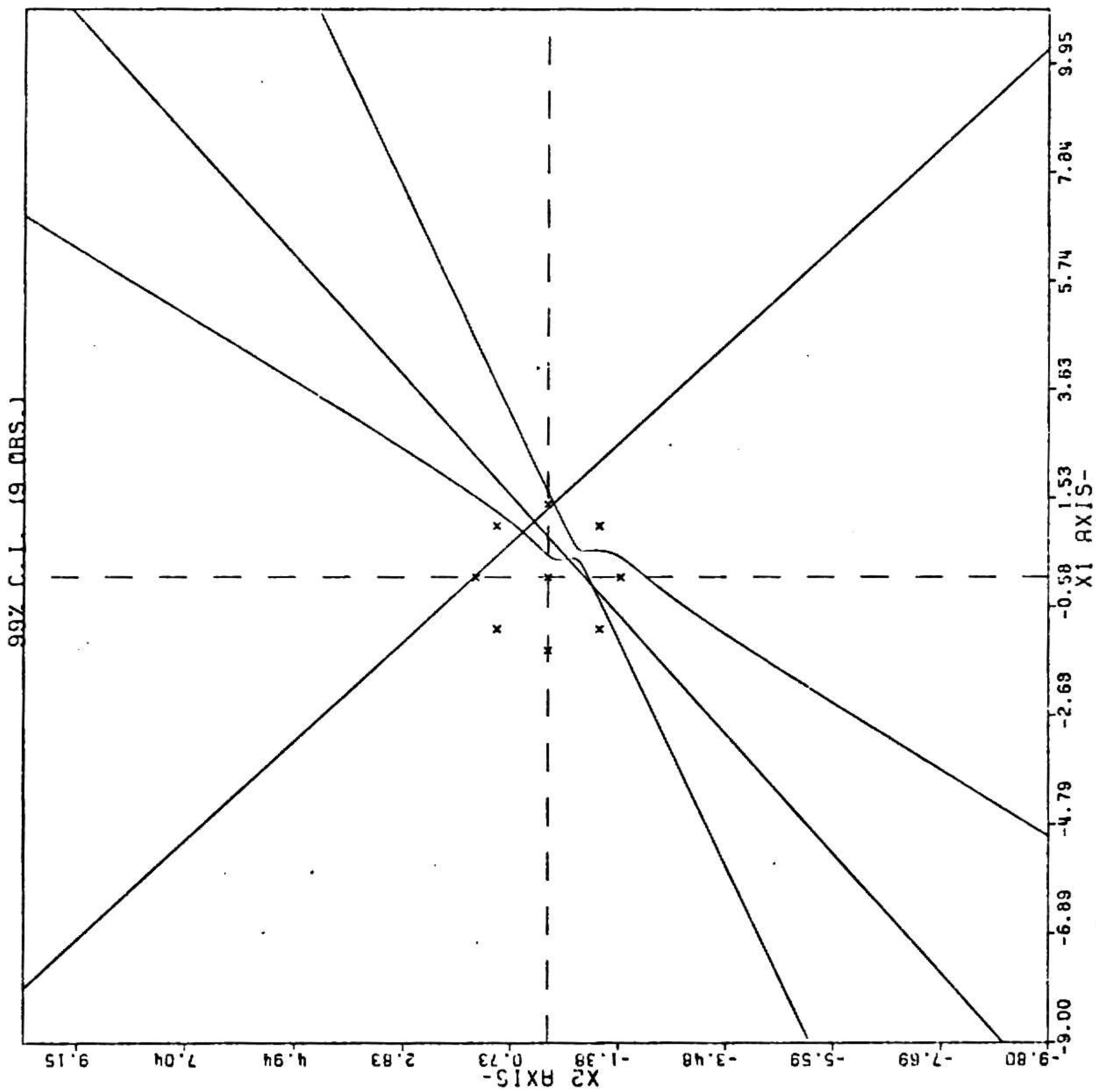
The parameter cards used are shown below and the plots follow in the next two pages.

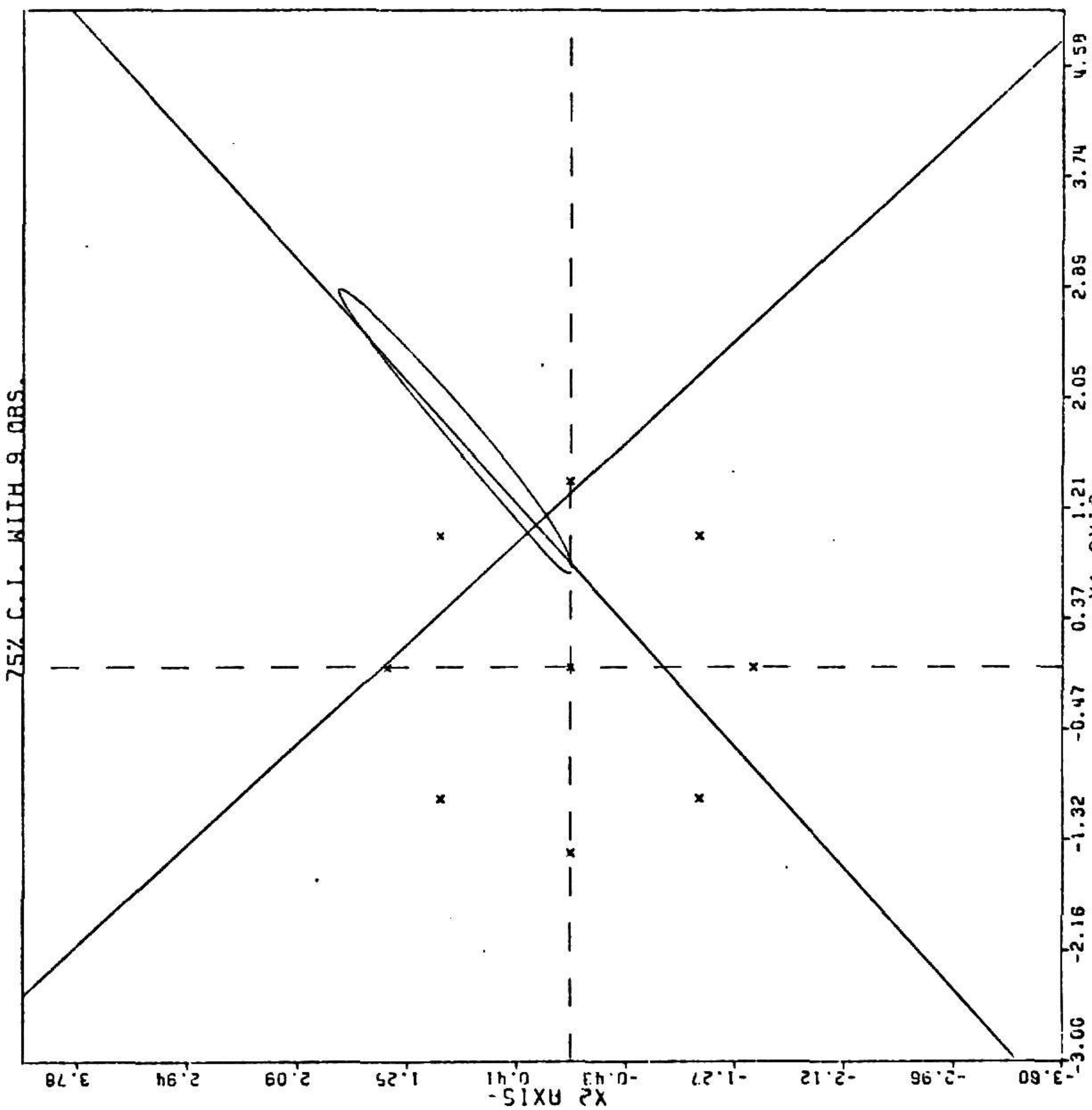
Example 1 b.

```
(3F10.3)            99% C.R. (9 OBS.)              -X1 AXIS-X2 AXIS-
  9 -1 17      9.5          30.8          -10.   10.  -10.   10.
```

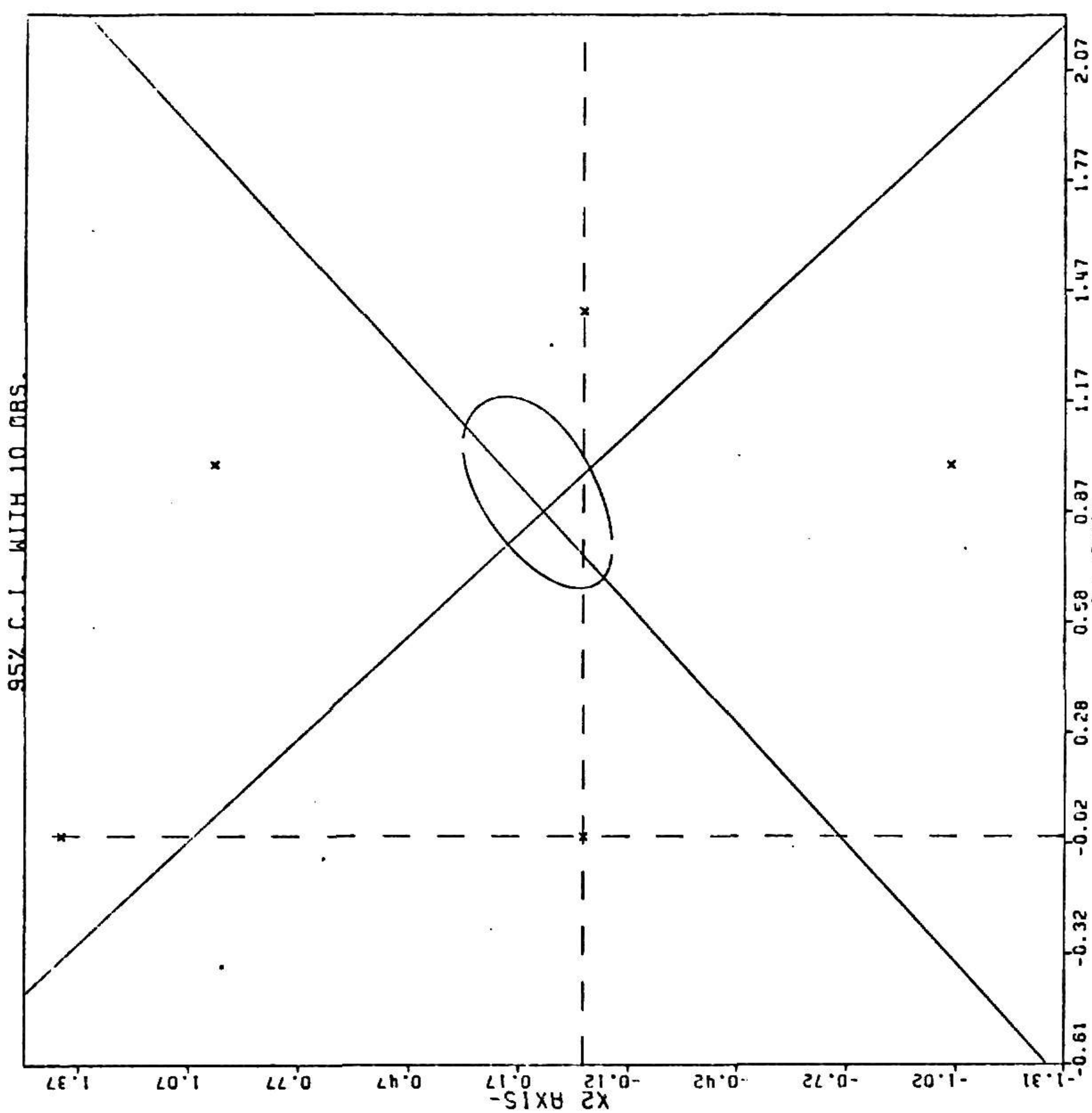
Example 1 c.

```
(3F10.3)            75% C.R. with 9 OBS.           -X1 AXIS-X2 AXIS-
  9 -1 20      9.5           2.28           -4.    4.   -4.    4.
```


99% C.I. (9 OBS.)
X1 AXIS-
9.95
7.84
5.74
3.63
1.53
-0.58
-2.69
-4.79
-6.89
-9.00
X2 AXIS-
9.15
7.04
4.94
2.83
0.73
-1.38
-3.48
-5.59
-7.69
-9.80

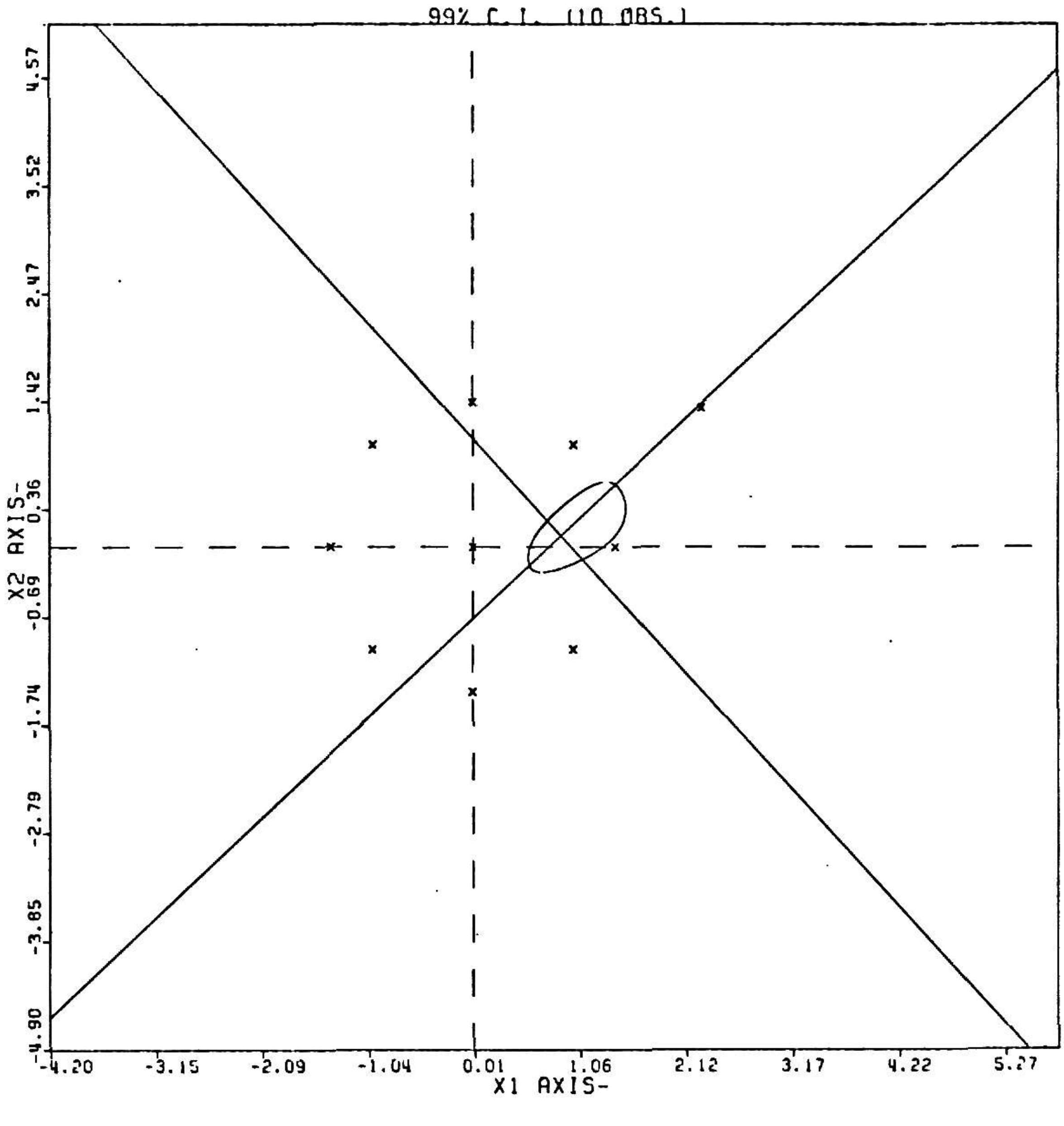

75% C.I. WITH 9 OBS.
X1 AXIS
-3.00 -2.16 -1.32 -0.47 0.37 1.21 2.05 2.89 3.74 4.59
X2 AXIS
-3.80 -2.96 -2.12 -1.27 -0.43 0.41 1.25 2.09 2.94 3.78

Example 1 d: This example demonstrates the effect an added observation has on the shape of the confidence region. Using the plot from Example 1 a., a tenth observation was generated along the major axis of the ellipsoidal confidence region. At the point (2.265, 1.354) the observation was 75.634, FVAL changed to $F_{.05}(2,4) = 6.94$ and ks^2F became 4.86755. The stationary point shifted slightly to (.87429, .10922). The plot on the next page shows that the confidence region closed significantly.

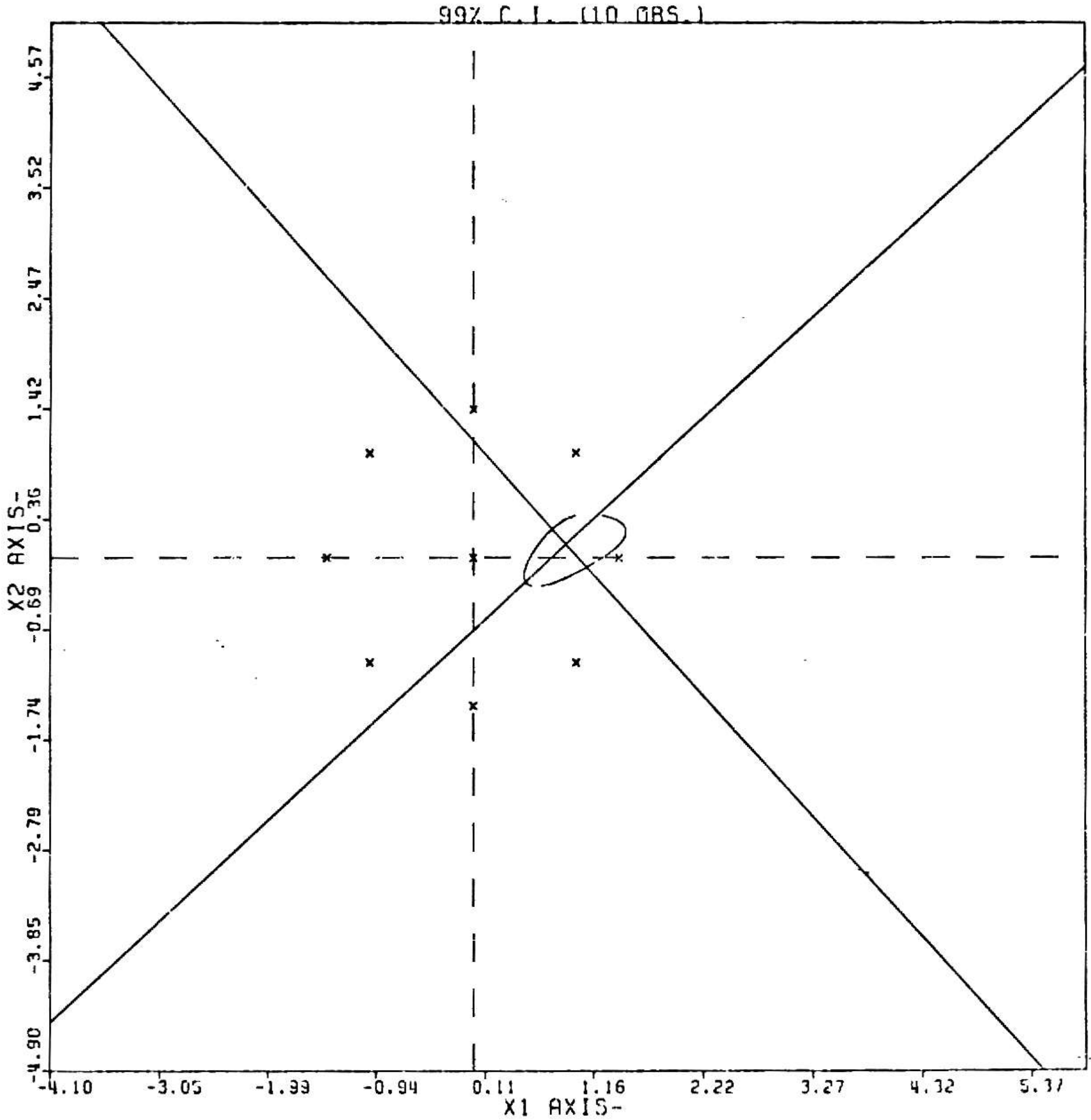

95% C.I. WITH 10 OBS.
X1 AXIS-
-0.61
-0.32
-0.02
0.28
0.58
0.87
1.17
1.47
1.77
2.07
X2 AXIS-
1.37
1.07
0.77
0.47
0.17
-0.12
-0.42
-0.72
-1.02
-1.31

Examples 1 e and 1 f: Repeating the idea used in Example 1 d., a tenth observation was added for the analysis of the 99% confidence region used in Example 1 b. On the first run the same tenth observation was used as in the previous example. On the next run the tenth observation was taken further from the origin of the canonical axes at (9.95, 8.492) with a response of -124.539. Comparing the plots on the next two pages, one can see that the different choices of tenth observation made little difference. The confidence region closed to about the same size with slightly different shapes for the two runs.

Other results given for Example 1 e. and 1 f. respectively are stationary points (.87429, .10922) and (.90270, .13012), a Turing's conditioning measure of 1.209901 compared to 1.214495, and $ks^2F = 12.62$ compared to 11.72.



99% C.I. (10 OBS.)
X2 AXIS-
4.57
3.52
2.47
1.42
0.36
-0.69
-1.74
-2.79
-3.85
-4.90
-4.20
-3.15
-2.09
-1.04
0.01
1.06
2.12
3.17
4.22
5.27
X1 AXIS-


99% C.I. (10 OBS.)
X2 AXIS-
4.57
3.52
2.47
1.42
0.36
-0.69
-1.74
-2.79
-3.85
-4.90
-4.10
-3.05
-1.99
-0.94
0.11
1.16
2.22
3.27
4.32
5.37
X1 AXIS-

Example 2: In this example the assumed response surface remains

$$\eta = 78.373 + 4.533x_1 - 1.867x_2 - 3.333x_1^2 - 3.333x_2^2 + 4.0x_1x_2.$$

To the nine observations of Example 1. are added six observations taken at the origin.

x_1	x_2	y
0	0	78.973
0	0	77.073
0	0	78.043
0	0	78.374
0	0	80.175
0	0	79.277

The source of variation table is now

Source	dof
Mean	1
Linear	2
Quadratic	3
Residual	9
Lack of Fit	3
Pure Error	6

The program has an option to perform a lack of fit test when data includes subsampling. IDUP = 0 tells the program to perform the lack of fit test using FLOFT = F(3,6) as the critical value. If the test is significant FVAL = F(2,6) will be the critical value of the confidence region, but if the test is not significant then FPOOL = F(2,9) will be used.

On the next page is an echo check of the input parameter cards. All the F-values are at the .05 level. The results of the Lack of Fit test are displayed below the echo check.

ECHO CHECK OF PARAMETER CARDS

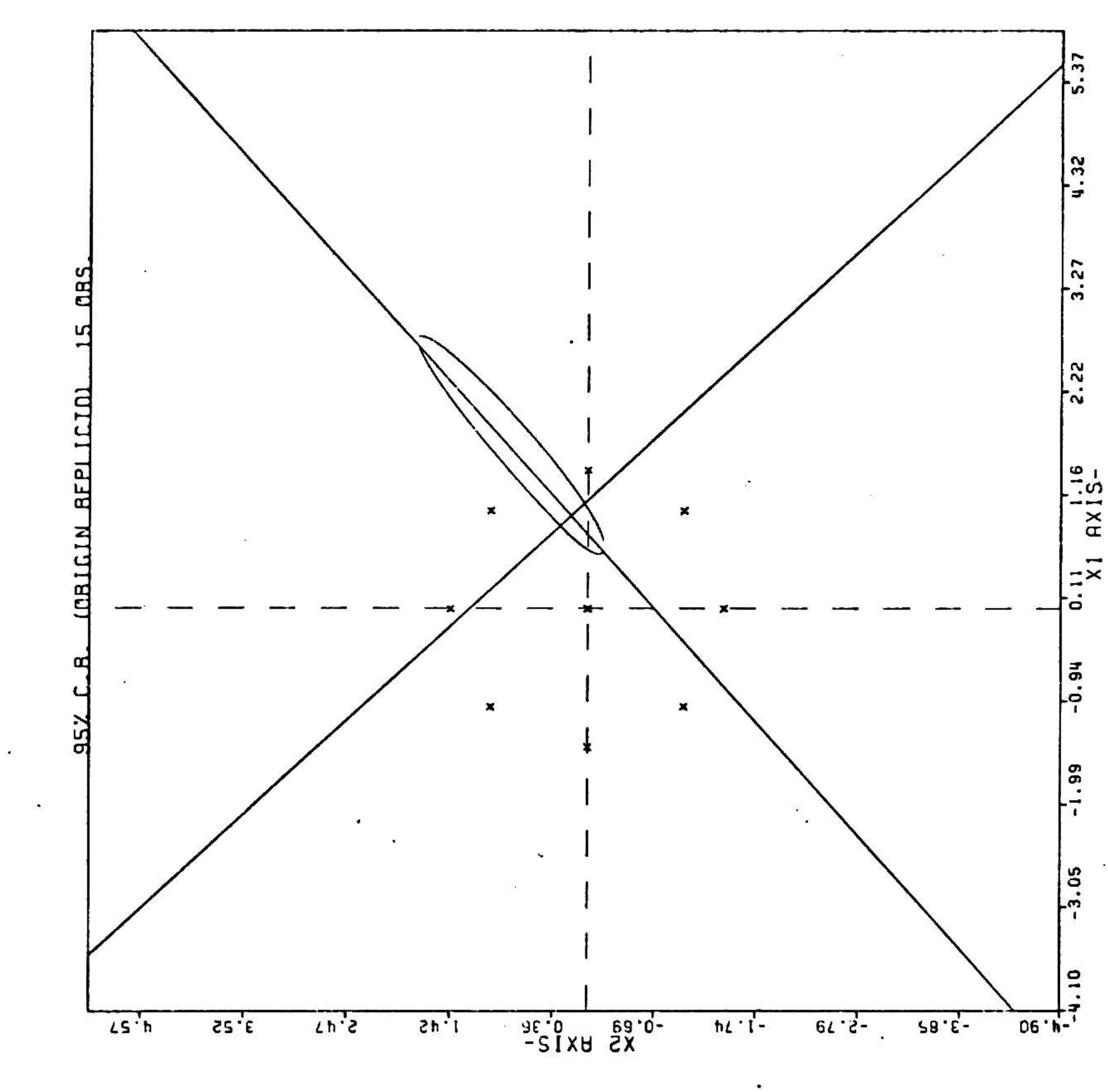
N	IDUP	NCHAR	PLTH	FLOFT	FVAL	FPOOL	X1-AXIS PLOT LIMITS		X2-AXIS PLOT LIMITS	
15	0	35	9.500	4.760	5.140	4.260	-5.000	5.000	-5.000	5.000

DATA INPUT FRMT	PLOT TITLE	X1AXIS LBL	X2AXIS LBL
(3F10.3)	95% C.R. (ORIGIN REPLICTD) 15 OBS. -	X1 AXIS-	X2 AXIS-

LACK OF FIT TEST IS NOT SIGNIFICANT. POOLED MEAN SQUARE IS USED.

The next page gives the results of the regression and canonical analyses. The second page following is the plot of the 95% confidence region with 15 observations.

```
                          95% C.R. (ORIGIN REPLICTD)  15 OBS.        -

        ERROR VARIANCE USED IS    0.777749     TO GIVE K*F*S**2 =     6.62643

REGRESSION COEFFICIENTS    0           1            2            11           22           12

                       78.5816      4.89289     -2.32673     -2.91736     -3.26347     3.64150

INVERSE MATRIX          1            2            11             22           12

              1   0.125019
              2   0.0          0.125019
             11   0.0          0.0          0.129522
             22   0.0          0.0          0.444674D-02   0.129522
             12   0.0          0.0          0.0            0.0          0.250000

TURINGS CONDITIONING MEASURE  C =    1.240581

STATIONARY POINT      EIGENVALUES     EIGENVECTORS

X(1)=   0.54523        -1.2615          0.73980  0.67282

X(2)=   0.17088        -4.9194         -0.67282  0.73980
```


95% C.R. (ORIGIN REPLICID) 15 OBS.
X1 AXIS-
-4.10
-3.05
-1.99
-0.94
0.11
1.16
2.22
3.27
4.32
5.37
X2 AXIS-
-4.90
-3.85
-2.79
-1.74
-0.69
0.36
1.42
2.47
3.52
4.57

CONCLUSION

In this paper, the three parts of a response surface analysis of a second-order response model have been discussed. Computational techniques used for the regression analysis, canonical analysis, and the determination of a confidence region for the stationary point have been shown.

A FORTRAN IV program which uses these computational techniques produces plots of the $(1-\alpha)100\%$ confidence region about the stationary point of the response surface. Several examples using this program have been shown.

Again the point should be made that these confidence regions are for the location of the stationary point $\underline{x}_0$ not for the response at $\underline{x}_0$. However, to some extent the confidence region does reflect what is happening in the contour system of the response surface.

Sometimes during a response surface analysis the researcher finds that the stationary point is remote from his design region. In this case the plots of the confidence region for the stationary point of the response surface should be of value when he is choosing experiment points for subsequent analyses.

REFERENCES

{1} Box, G. E. P. and J. S. Hunter (1954). "A Confidence Region for the Solution of a Set of Simultaneous Equations with an Application to Experimental Design". Biometrika, 41: 190-199.

{2} Design and Analysis of Industrial Experiments. editor O. L. Davies (1954). Hafner Publishing Company, New York.

{3} Draper, N. R. and H. Smith (1966). Applied Regression Analysis. John Wiley & Sons, Inc., New York.

{4} Meyers, Raymond H. (1971). Response Surface Methodology. Allyn and Bacon, Inc., Boston.

{5} Morrison, Donald F. (1967). Multivariate Statistical Methods. McGraw-Hill, New York.

A P P E N D I X

USER'S GUIDE

Confidence Region Plotting Package

(Second-Order Model with 2 Independent Variables)

Main Program: Carries out options on the desired error variance, estimates the coefficients of the response function, performs the canonical analysis, and plots the confidence region for the stationary point of a second-order polynomial model with 2 independent variables. CALCOMP plot - FORTRAN IV, G LEVEL

Subroutines:

RGRES -- Performs a regression analysis of Y on X1, X2, returning $(X'X)^{-1}$, the estimated regression coefficients, and sums of squares due to residuals or due to pure error, if there is subsampling. (double precision)

FINR -- Computes elements of the equation for the confidence region boundary (double precision)

POLYS -- Computes the roots of a polynomial, returns the real roots, and returns the number of real roots (double precision)

NOTE: Since CALCOMP plotting routines are single precision, any scanning of arrays should use an increment of 2.

Input Description:

Data Card 1 -- Cols. 1-3 N=no. of observations

Cols. 4-6 IDUP=-1 if Residual MS is used as error estimate
0 if LOF test is to be performed
1 if Pure Error MS is to be used (prior knowledge of results of LOF test)

Cols. 7-9 NCHAR=no. of characters in title of plot (max = 43)

Cols. 10-15 PLTH=length of vertical (X2) axis (max = 9.5 inches)

Cols. 16-21 FLOFT=critical F value for LOF test

Cols. 22-27 FVAL =critical F value for confidence region if IDUP not 0

=critical F value for confidence region based on Pure Error MS if IDUP = 0

Cols. 28-33 FPOOL=critical F value for confidence region based on pooled MS

Cols. 34-39 VL(1)=lower limit of plot for X1

Cols. 40-45 VU(1)=max X1

Cols. 46-51 VL(2)=min X2

Cols. 52-57 VU(2)=max X2

Data Card 2 -- Cols. 1-16 data format for inputting X1, X2, and Y any acceptable FORTRAN format allowed

Cols. 17-59 characters of plot title

Col. 60 underscore

Cols. 61-67 X1 (horizontal) axis label

Col. 68 underscore

Cols. 69-75 X2 (vertical) axis label

Col. 76 underscore

Data Card 3 thru N + 2 Contain the input values of X1, X2, and Y punched in the format specified on Data Card 2, Cols. 1-16

Data Card N + 3 Blank card if the program is to be ended

Any number of plots may be run but the set of data cards for the last plot must be followed by a blank card.

Output Description:

A. Print-out

1. Title
2. Error variance used in the confidence region and the value ks^2F obtained
3. Estimated Regression Coefficients
4. $(X'X)^{-1}$
5. Stationary Point, Characteristic roots and vectors

B. Plot

The entire plot is approximately centered on the estimated stationary point. Both the X1,X2 axes and the canonical axes are plotted. Each observation point within the range of the plot is also marked. The Confidence Region (CR) is plotted continuously. The range of the horizontal axis is VU(1) - VL(1) but will not be centered on VU(1)/2 - VL(1)/2 due to centering on the estimated stationary point. The same is true for the vertical axis.
The length of the horizontal axis is set equal to the length of the vertical axis in the program. This is done partly to aid in making the canonical axis plot orthogonally. Thus equal scaling must be chosen on each axis to obtain orthogonal canonical axes.

NOTES:

If IDUP = -1, then only FVAL needs to be specified

= +1, then only FVal needs to be specified

= 0, then FLOFT, FVAL, FPOOL must all be specified

Execution time should run under 20 seconds for most plots. If this is exceeded then the range of the plot may be too wide for the program to handle.

```
                C
                C                         CONFIDENCE REGION PLOTTING PACKAGE
                C
                C                   PLOTS CONFIDENCE REGION FOR STAIONARY POINT
                C                   OF A SECOND ORDER RESPONSE SURFACE
                C
                C
0001                 IMPLICIT REAL*8(A-H,O-Z)
0002                 REAL PTW,YORG,PLTW,X1LB,X2LB,HEIG,PTICW,PTICH,VL,VU,SSYM,PLTH,Z1,
                    1Z2,XX,XY,ZT1,ZT2,H
0003                 DIMENSION BB(6),CC(5,5),E(6,6),R(5,15),P(5,5),V(15),X(5),AM(2,2),
                    1X1(30),X2(30),Y(30),XL(2),XU(2),XS(2),VL(2),VU(2),IFM(6),XY(600),
                    2H(11),LV(6),AL(2),XX(2400),TPS(4),X1LB(2),X2LB(2)
0004                 DIMENSION IBUF(1000)
0005                 COMMON A,B,C,D
                C
                C
                C    N=NO. OF OBSERVATIONS
                C
                C          -1 IF RESIDUAL MEAN SQUARE IS USED
                C    IDUP= 0 IF LACK OF FIT TEST IS TO BE PERFORMED
                C          1 IF DUPLICATION MEAN SQUARE IS USED
                C
                C     NCHAR=NO. OF CHARACTERS TO BE DRAWN IN TITLE OF PLOT (LIMIT OF 43)
                C     PLTH=LENGTH OF X(2)-AXIS IN INCHES (LESS THAN 9.5)
                C    FLOFT=CRITICAL F VALUE FOR LACK OF FIT TEST
                C    FVAL=CRITICAL F VALUE FOR CONFIDENCE REGION, IF IDUP NOT 0.
                C    FVAL=CRITICAL F VALUE FOR CONFIDENCE REGION BASED ON DUPLICATION
                C             MEAN SQUARE, IF IDUP=0.
                C    FPOOL=CRITICAL F VALUE FOR CONFIDENCE REGION BASED ON POOLED MEAN
                C          SQUARES.
                C
                C     VL(I)=LOWER LIMIT OF PLOT FOR X(I)
                C     VU(I)=UPPER LIMIT OF PLOT FOR X(I)
                C
                C    IFM=VECTOR CONTAINING FORMAT FOR INPUT OF OBSERVATIONS
                C     H=VECTOR CONTAINING THE PLOT TITLE. THE CHARACTER IN THE
                C             (NCHAR+1)-POSITION MUST BE THE UNDERSCORE CHARACTER _.
                C     X1LB=VECTOR CONTAINING THE X(1)-AXIS LABEL.
                C     X2LB=VECTOR CONTAINING THE X(2)-AXIS LABEL.
                C
                C          X1LB AND X2LB CAN HAVE 1-8 CHARACTERS EACH, BUT THE
                C          LAST CHARACTER MUST BE THE UNDERSCORE _.
                C
                C
0006                 CALL PLOTS(IBUF,1000)
0007                 KOUNT=0
0008             100 READ(1,40)N,IDUP,NCHAR,PLTH,FLOFT,FVAL,FPOOL,(VL(I),VU(I),I=1,2)
0009              40 FORMAT(3I3,2F6.0)
0010                 WRITE(6,45)
0011              45 FORMAT('1ECHO CHECK OF PARAMETER CARDS'/'0 N',3X,'IDUP',3X,'NCHAR'
                    $,3X,'PLTH',4X,'FLOFT',4X,'FVAL',5X,'FPOOL',5X,'X1-AXIS PLOT LIMITS
                    $',4X,'X2-AXIS PLOT LIMITS')
0012                 WRITE(6,44)N,IDUP,NCHAR,PLTH,FLOFT,FVAL,FPOOL,(VL(I),VU(I),I=1,2)
0013              44 FORMAT(1H0,I3,3X,I3,4X,I3,4F9.3,5X,2F9.3,5X,2F9.3,///)
0014                 IF(N)587,987,141
0015             141 READ(1,47)(IFM(J),J=1,4),(H(I),I=1,11),X1LB(1),X1LB(2),X2LB(1),
                    1X2LB(2)
```

```
0016         42 FORMAT(19A4)
0017            WRITE(6,48)
0018         48 FORMAT(1H0,'DATA INPUT FRMT',15X,'PLOT TITLE',24X,'X1AXIS LBL',5X,
               $'X2AXIS LBL')
0019            WRITE(6,47)(IFM(J),J=1,4),(H(I),I=1,11),X1LB(1),X1LB(2),X2LB(1),X2
               $LB(2)
0020         47 FORMAT(1H0,15A4,5X,2A4,5X,2A4////)
0021            DO 41 I=1,N
0022            READ(1,IFM)X1(I),X2(I),Y(I)
0023         41 CONTINUE
0024            DO 804 I=1,3
0025            LV(I)=I-1
0026        804 LV(I+3)=I+I*10
0027            LV(6)=12
0028            ZEP=10.D-10
          C
          C     * COMPUTE REGRESSION COEFFICIENTS AND X'X INVERSE
          C
0029            CALL RGRES(N,X1,X2,Y,BB,CC,S2,SLOF,IDFS,ZER)
0030            IF(IDFS)701,701,700
0031        701 IF(IDUP)700,702,702
0032        702 WRITE(6,1000)
0033       1000 FORMAT('0SINCE THERE IS NO SUBSAMPLING, A LACK OF FIT TEST IS INAP
               $PROPRIATE.',/,' CHECK IDUP AND FVAL VALUES.')
0034            GO TO 100
0035        700 IF(IDUP)25,26,27
          C
          C     * DETERMINE ERROR VARIANCE FOR CONFIDENCE REGION
          C
0036         25 S=SLOF/(N-6-IDFS)
0037            GO TO 28
0038         27 S=S2/IDFS
0039            GO TO 28
0040         26 SLOF=SLOF/FLOAT(N-6-IDFS)
0041            S2=S2/FLOAT(IDFS)
0042            IF(SLOF/S2-FLOFT)30,29,29
0043         29 S=S2
0044            WRITE(6,1001)
0045       1001 FORMAT(1H0,40X,'LACK OF FIT TEST IS SIGNIFICANT')
0046            GO TO 28
0047         30 FVAL=FPCFL
0048            WRITE(6,1002)
0049       1002 FORMAT(1H0,40X,'LACK OF FIT TEST IS NOT SIGNIFICANT. POOLED MEAN
               $SQUARE IS USED.')
0050            S=(IDFS*S2+(N-6-IDFS)*SLOF)/(N-6)
0051         28 CON=2.*FVAL*S
0052            WRITE(6,203)(H(I),I=1,11)
0053        203 FORMAT('1',/'0',40X,11A4//)
0054            WRITE(6,60) S,CON
0055         60 FORMAT(10X,'ERROR VARIANCE USED IS',G16.6,' TO GIVE K*F*S**2 =',
               $G16.6//)
0056            WRITE(6,806)(LV(I),I=1,6)
0057            WRITE(6,807)(BB(I),I=1,6)
0058            WRITE(6,808)(LV(I),I=2,6)
0059            DO 809 I=1,5
0060            LV(I)=LV(I+1)
0061        809 WRITE(6,810)LV(I),(CC(I,J),J=1,I)
0062        806 FORMAT('0REGRESSION COEFFICIENTS',6(I5,10X)/)
```

```
  808 FORMAT(' INVERSE MATRIX',1X,5(I8,8X)/)
  807 FORMAT(20X,6G15.6//)
  810 FORMAT(10X,I4,5G16.6/)
C
C    * COMPUTE QUADRATIC DELTA TERMS AND VARIANCE-COVARIANCE MATRIX
C        OF DELTAS IN TERMS OF COEFFICIENTS OF POLYNOMIALS IN
C        PSI(1) AND PSI(2)
C
       DO 1I=1,5
      DO 1J=1,5
    1 P(I,J)=BB(I+1)*BB(J+1)
      ICN=0
      I=1
    2 J=I+1
      K=J+1
      E(I,1)=P(1,1)
      E(J,1)=P(2,2)
      E(K,1)=P(1,2)
      E(I,2)=P(1,3)*4.
      E(J,2)=P(2,5)*2.
      E(K,2)=P(2,3)*2.+P(1,5)
      E(I,3)=P(1,5)*2.
      E(J,3)=P(2,4)*4.
      E(K,3)=P(1,4)*2.+P(2,5)
      E(I,4)=P(3,3)*4.
      E(J,4)=P(5,5)
      E(K,4)=P(3,5)*2.
      E(I,5)=P(5,5)
      E(J,5)=P(4,4)*4.
      E(K,5)=P(4,5)*2.
      E(I,6)=P(3,5)*4.
      E(J,6)=P(4,5)*4.
      E(K,6)=P(3,4)*4.+P(5,5)
      IF(ICN)5,5,4
    5 DO 3 I=1,5
      DO 3 J=1,5
    3 P(I,J)=CC(I,J)
      I=4
      ICN=1
      GO TO 2
C
C    * COMPUTE COEFFICIENTS FOR PSI(1) AND PSI(2) IN QUARTIC
C        POLYNOMIAL REPRESENTING CONFIDENCE REGION BOUNDARY
C
    4 CALL FINR(V,1,5,E)
      DO   7 J=1,15
    7 R(1,J)=V(J)
      CALL FINR(V,2,4,E)
      DO   8 J=1,15
    8 R(2,J)=V(J)
      CALL FINR(V,3,6,E)
      DO   9 J=1,15
    9 R(3,J)=-2.*V(J)
      CALL FINR(V,4,5,E)
      DO  10 J=1,15
   10 R(4,J)=-CON*V(J)
      CALL FINR(V,6,6,E)
      DO  11 J=1,15
```

```
0112           V(J)=CCN*V(J)
0113           R(5,J)=V(J)
0114           DO 11 I=1,4
0115        11 V(J)=V(J)+R(I,J)
          C
          C    * DETERMINE SIZE OF DESIGN REGION
          C
0116           CO 601 I=1,2
0117           XU(I)=-1000.
0118       601 XL(I)= 1000.
0119           CC 117 I=1,N
0120           IF(XU(1)-X1(I))112,112,111
0121       112 XU(1)=X1(I)
0122       111 IF(XL(1)-X1(I))113,114,114
0123       114 XL(1)=X1(I)
0124       113 IF(XU(2)-X2(I))115,115,116
0125       115 XU(2)=X2(I)
0126       116 IF(XL(2)-X2(I))117,110,110
0127       110 XL(2)=X2(I)
0128       117 CCNTINUE
          C
          C    * DETERMINE STATIONARY PCINT AND ITS LOCATION RELATIVE TO
          C         DESIGN REGICN
          C
0129           NC=1
0130           DET=-2.*BB(4)*BB(5)+.5*BB(6)*BB(6)
0131           IF(DABS(DET)-ZER)118,118,119
0132       118 NC=0
0133           GC TC 120
0134       119 XS(1)=(BB(5)*BB(2)-.5*BB(6)*BB(3))/DET
0135           XS(2)=(BB(4)*BB(3)-.5*BB(6)*BB(2))/DET
          C
          C    * COMPUTE EIGENVALUES ANC EIGENVECTORS FOR CANONICAL
          C         TRANSFCRMATICN
          C
0136       624 TEM=BB(4)+BB(5)
0137           TE2=DSCRT(TEM*TEM+2.*DET)
0138           AL(1)=.5*(TEM+TE2)
0139           AL(2)=.5*(TEM-TE2)
0140           AT=BB(4)-AL(1)+.5*BB(6)
0141           BT=BB(5)-AL(1)+.5*BB(6)
0142           AM(1,1)=DSCRT(1./(1.+(AT/BT)**2))
0143           AM(2,1)=-AT*AM(1,1)/BT
0144           AM(2,2)=AM(1,1)
0145           AM(1,2)=-AM(2,1)
0146           COND=C.5*DSQRT((AL(1)+AL(2))*((1./AL(1))+(1./AL(2))))
          C    * PRINT RESULTS OF CANONICAL ANALYSIS
0147           WRITE(6,811)COND
0148       811 FORMAT(//' TURINGS CONDITIONING MEASURE  C =',F12.6/)
0149       801 WRITE(6,802)
0150       802 FORMAT('OSTATICNARY POINT',7X,'EIGENVALUES',6X,'EIGENVECTORS'//)
0151           CC 904 I=1,2
0152       904 WRITE(6,803)I,XS(I),AL(I),AM(1,I),AM(2,I)
0153       803 FORMAT(' X(',I1,')=',G14.5,G16.5,5X,2F9.5/)
0154           IF(XS(1)-XL(1))120,120,121
0155       121 IF(XS(1)-XU(1))122,120,120
0156       122 IF(XS(2)-XL(2))120,12C,123
0157       123 IF(XS(2)-XU(2))124,120,120
```

```
0158      124 DC 127 I=1,2
0159          IXS=XS(I)*10.
0160          VU(I)=.1*IXS+VU(I)
0161      127 VL(I)=.1*IXS+VL(I)
        C
        C    * DETERMINE SCALES, DRAW AXES AND LABELS
        C
0162      120 HEIG=-PLTH
0163          SCAL=PLTH/(VU(2)-VL(2))
0164          PLTW=(VU(1)-VL(1))*SCAL
0165          SSYM=.5*PLTW-.0643*NCHAR
0166          IF(SSYM+.5)510,511,511
0167      510 SSYM=-.5
0168      511 KOUNT=KOUNT+1
0169          CALL PLOT(0.0,-11.,3)
0170          CALL PLOT(0.,-10.,-3)
0171          SCLE1=(VU(1)-VL(1))/PLTW
0172          SCLE2=(VU(2)-VL(2))/PLTH
0173          IF(KOUNT.EQ.1)GO TO 512
0174          CALL PLOT(12.0,0.0,-3)
0175      512 CALL AXIS(0.0,0.,X1LB,-8,PLTW,0.,VL(1),SCLE1)
0176          CALL AXIS(0.,0.,X2LB,8,PLTH,90.,VL(2),SCLE2)
0177          HEIG=PLTH+.1
0178          CALL SYMBOL(SSYM,PLTH,.14,H,0.,NCHAR)
0179          CALL PLOT(PLTW,PLTH,3)
0180          CALL PLOT(0.,PLTH,2)
0181          DO 400 I=1,N
0182          IF(X1(I)-VL(1))400,400,402
0183      402 IF(X1(I)-VU(1))403,400,400
0184      403 IF(X2(I)-VL(2))400,400,404
0185      404 IF(X2(I)-VU(2))405,400,400
0186      405 Z1=(X1(I)-VL(1))*SCAL
0187          Z2=(X2(I)-VL(2))*SCAL
0188          CALL PLOT(Z1,Z2,3)
0189          CALL SYMBOL(Z1,Z2,.07,4,0.,-1)
0190      400 CONTINUE
        C
        C    * DRAW CANONICAL AXES
        C
0191          IF(NC)501,501,502
0192      502 DO 420 J=1,2
0193          INUM=3
0194          DO 406 M=1,2
0195          IF(M-1)407,407,408
0196      407 K=2
0197          GO TO 409
0198      408 K=1
0199      409 DO 406 L=1,2
0200          IF (L-1)411,411,412
0201      411 TEV=VL(K)
0202          GO TO 413
0203      412 TEV=VU(K)
0204      413 ZT=XS(M)-AM(K,J)*(TEV-XS(K))/AM(M,J)
0205          IF(ZT-VL(M))406,414,414
0206      414 IF(ZT-VU(M))415,415,406
0207      415 IF(M-1)416,416,417
0208      416 Z1=(ZT-VL(1))*SCAL
0209          Z2=(TEV-VL(2))*SCAL
```

```
0210           GC TO 418
0211       417 Z1=(TEV-VL(1))*SCAL
0212           Z2=(ZT-VL(2))*SCAL
0213       418 CALL PLOT(Z1,Z2,INUM)
0214           IF(INUM-2)420,420,410
0215       410 INUM=2
0216       406 CONTINUE
0217       420 CONTINUE
0219       501 INUM=2
0219           M=0
0220           DH=(VU(2)-VL(2))*.02/PLTH
0221           Z=VL(2)-DH
0222       422 I=C
0223       423 Z=Z+DH
0224           IF(Z-VU(2))424,424,425
0225       425 IF(M)500,500,426
0226       426 INUM=3
0227           GO TO 450
         C
         C     * DETERMINE AND SCLVE QUARTIC POLYNOMIAL IN PSI(1) FOR FIXED PSI(2)
         C
0228       424 TEM= V(11)
0229           IF(DABS(TEM)-ZER)101,101,102
0230       101 WRITE(6,200)
0231       200 FORMAT('1QUARTIC EQUATICN REDUCES TO CUBIC')
0232           GO TO 100
0233       102 TE2=Z*Z
0234           TE3=Z*TE2
0235           A=(V(7)+Z*V(13))/TEM
0236           B=(V(4)+Z*V(9)+TE2*V(14))/TEM
0237           C=(V(2)+Z*V(6)+TE2*V(10)+TE3*V(15))/TEM
0238           D=(V(1)+Z*V(3)+TE2*V(5)+TE3*V(8)+Z*TE3*V(12))/TEM
0239           IDEG=5
0240           CALL POLYS(X,NR,IDEG,ZER)
0241           K=0
         C
         C     * DETERMINE NO. OF REAL ROOTS IN RANGE OF PLOT
         C
0242           IF(NR)427,427,428
0243       428 DO 429 J=1,NR
0244           IF(X(J)-VU(1))430,430,429
0245       430 IF(X(J)-VL(1))429,431,431
0246       431 K=K+1
0247           TPS(K)=X(J)
0248       429 CONTINUE
0249           IF(K)427,427,432
0250       427 IF(M)423,423,450
0251       450 NPT=I
0252           GC TO 475
0253       432 I=I+1
0254           IF(I-1)433,433,434
0255       433 M=K
         C
         C     * PLOT THE CONFIDENCE REGION BOUNDARY
         C
0256       439 XY(I)=(Z-VL(2))*SCAL
0257           DC 435 J=1,K
0258           ZT=VU(1)
```

```
0259            MM=(J-1)*600+I
0260            DO 436 L=1,K
0261            IF(TPS(L)-ZT)437,437,436
0262        437 ZT=TPS(L)
0263            LL=L
0264        436 CONTINUE
0265            XX(MM)=(ZT-VL(1))*SCAL
0266        435 TPS(LL)=VU(1)+1.
0267            GO TO 423
0268        434 IF(K-M)438,439,438
0269        438 NPT=I-1
0270        475 DO 479 J=1,M
0271            L=(J-1)*600+1
0272        476 IND=(L+NPT)-1
0273            INC=1
0274            IPIN=3
0275            DO 503 IN=L,IND
0276            CALL PLOT(XX(IN),XY(INC),IPIN)
0277            IPIN=2
0278            INC=INC+1
0279        503 CONTINUE
0280        479 CONTINUE
0281            IF(INUM-2)477,477,500
0282        477 M=K
0283            IF(K)422,422,478
0284        478 I=1
0285            GO TO 439
0286        500 CALL PLOT(PLTW,PLTH,3)
0287            CALL PLOT(PLTW,0.,2)
          C
          C     * DRAW ORIGINAL AXES
          C
0288            DO 461 J=1,2
0289            IF(VL(J))460,461,461
0290        460 IF(VU(J))461,461,463
0291        463 Z1=-VL(J)*SCAL
0292            ZT1=-.5
0293        470 ZT1=ZT1+.5
0294            ZT2=ZT1+.25
0295            IF(J-1)462,462,464
0296        462 IF(ZT2-PLTH)465,465,461
0297        465 CALL PLOT(Z1,ZT1,3)
0298            CALL PLOT(Z1,ZT2,2)
0299            GO TO 470
0300        464 IF(ZT2-PLTW)466,466,461
0301        466 CALL PLOT(ZT1,Z1,3)
0302            CALL PLOT(ZT2,Z1,2)
0303            GO TO 470
0304        461 CONTINUE
0305            GO TO 100
0306        987 CALL PLOT(0.,0.,999)
0307            STOP
0308            END
```

```
FORTRAN IV G LEVEL  21              RGRES              DATE = 75205          12/48/58                 PAGE 0001

0001                  SUBROUTINE RGRES(NOBS,X1,X2,Y,B,C,S,XLOF,IDFS,ZER)
0002                  IMPLICIT REAL*8(A-H,O-Z)
0003                  DIMENSION X1(30),X2(30),Y(30),B(6),A(7,14),ISW(25),C(5,5),X(7)
0004                  COMMON R,T,U,V
            C
            C         ZERO OUT SUMMATION AREAS AND FORM IDENTITY MATRIX
            C
0005                  IDFS=0
0006                  S=0.0
0007                  DO 10 I=1,7
0008                  DO 9 J=1,14
0009        9         A(I,J)=0.0
0010        10        A(I,I+7)=1.0
0011                  DO 12 I=1,NOBS
0012        12        ISW(I)=0
            C
            C         FORM X'X MATRIX
            C
0013                  DO 15 N=1,NOBS
0014                  X(1)=1.0
0015                  X(2)=X1(N)
0016                  X(3)=X2(N)
0017                  X(4)=X1(N)*X1(N)
0018                  X(5)=X2(N)*X2(N)
0019                  X(6)=X1(N)*X2(N)
0020                  X(7)=Y(N)
0021                  DO 15 I=1,7
0022                  DO 15 J=1,7
0023        15        A(I,J)=A(I,J)+X(I)*X(J)
            C
            C         ADJUST THE X'X MATRIX
            C
0024                  DO 35 J=1,6
0025                  D=A(J,J)
0026                  IF(D-ZER)34,34,25
0027        25        DO 30 K=1,14
0028        30        A(J,K)=A(J,K)/D
0029                  A(J,J)=1.0
0030                  DO 33 I=1,7
0031                  IF(I-J)31,33,31
0032        31        B1=A(I,J)
0033                  DO 32 K=1,14
0034        32        A(I,K)=A(I,K)-B1*A(J,K)
0035                  A(I,J)=0.0
0036        33        CONTINUE
0037                  GO TO 35
0038        34        DO 37 M=1,14
0039        37        A(J,M)=0.0
0040                  DO 36 I=1,7
0041        36        A(I,J)=0.0
0042                  A(J,J)=1.0
0043        35        CONTINUE
            C
            C         CHECK FOR INDEPENDENT VALUES OF X1 AND X2
            C
0044                  N=NOBS-1
0045                  DO 55 I=1,N
0046                  NDF=0
```

```
FORTRAN IV G LEVEL  21                 RGRES              DATE = 75205          12/48/58              PAGE 0002

 0047              IF(ISW(I)-1)40,55,40
 0048        40    J1=I+1
 0049              DO 50 J=J1,NOBS
 0050              IF(X1(I)-X1(J))50,44,50
 0051        44    IF(X2(I)-X2(J))50,46,50
 0052        46    IF(ISW(I)-1)47,48,47
 0053        47    SY2=Y(I)*Y(I)+Y(J)*Y(J)
 0054              SY=Y(I)+Y(J)
 0055                NDF=1
 0056              ISW(I)=1
 0057              ISW(J)=1
 0058              GO TO 50
 0059        48    SY2=SY2+Y(J)*Y(J)
 0060              SY=SY+Y(J)
 0061              ISW(J)=1
 0062              NDF=NDF+1
 0063        50    CONTINUE
 0064              IF(NDF)52,55,52
 0065        52    CSS=SY2-SY*SY/(NDF+1)
 0066              S=S+CSS
 0067              IOFS=IOFS+NDF
 0068        55    CONTINUE
             C
             C     FILL IN C MATRIX AND B VALUES AND COMPUTE LOF SS
             C
 0069              DO 60 I=1,5
 0070              DO 60 J=1,5
 0071        60    C(I,J)=A(I+1,J+8)
 0072              XLOF=A(7,7)-S
 0073              DO 62 I=1,6
 0074        62    B(I)=A(I,7)
 0075              RETURN
 0076              END
```

```
FORTRAN IV G LEVEL  21              FINR              DATE = 75205          12/48/58              PAGE 0001

              SUBROUTINE FINR(V,I,J,E)
              IMPLICIT REAL*8(A-H,O-Z)
              DIMENSION V(15),E(6,6),A(6),B(6)
              COMMON R,S,T,U
              DO 1 K=1,6
              A(K)=E(I,K)
            1 B(K)=E(J,K)
              V( 1)=A(1)*B(1)
              V(11)=A(4)*B(4)
              V(12)=A(5)*B(5)
              V( 2)=A(1)*B(2)+A(2)*B(1)
              V( 3)=A(1)*B(3)+A(3)*B(1)
              V( 7)=A(2)*B(4)+A(4)*B(2)
              V( 8)=A(3)*B(5)+A(5)*B(3)
              V(13)=A(4)*B(6)+A(6)*B(4)
              V(15)=A(5)*B(6)+A(6)*B(5)
              V( 4)=A(4)*B(1)+A(1)*B(4)+A(2)*B(2)
              V( 5)=A(5)*B(1)+A(1)*B(5)+A(3)*B(3)
              V(14)=A(4)*B(5)+A(5)*B(4)+A(6)*B(6)
              V( 6)=A(1)*B(6)+A(6)*B(1)+A(2)*B(3)+A(3)*B(2)
              V( 9)=A(2)*B(6)+A(6)*B(2)+A(3)*B(4)+A(4)*B(3)
              V(10)=A(2)*B(5)+A(5)*B(2)+A(3)*B(6)+A(6)*B(3)
              RETURN
              END
```

```
      SUBROUTINE PCLYS(X,NR,N,ZER)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION A(20),B(20),C(20),X(5)
      COMMON R,S,T,U
      M=0
      A(1)=1.
      A(2)=R
      A(3)=S
      A(4)=T
      A(5)=U
      NR=0
      B(1)=0.0
      C(1)=0.0
      C(2)=0.0
      B(2)=0.0
      P=1.0
      Q=1.0
    7 DO 4 I=1,N
      B(I+2)=A(I)+P*B(I+1)+Q*B(I)
    4 C(I+2)=B(2+I)+P*C(I+1)+Q*C(I)
      DEL=C(N)*C(N)-C(N+1)*C(N-1)
      IF(DABS(DEL)-ZER)6,6,5
    6 P=P+1.
      Q=Q+1.
      M=M+1
      IF(M-20)7,8,7
    5 DP=(B(N+2)*C(N-1)-B(N+1)*C(N))/DEL
      DQ=(B(N+1)*C(N+1)-C(N)*B(N+2))/DEL
      P=P+DP
      Q=Q+DQ
      ERR=DABS(DP)+DABS(DQ)
      IF(ERR-.000001)9,9,7
    9 DISC=P*P+4.*Q
      IF(DISC)11,10,10
   10 XAR=P/2+DSQRT(DISC)/2.
      XBR=P/2-DSQRT(DISC)/2.
      GO TO 31
   11 XAR=P/2.
      XBR=XAR
      XAI=DSQRT(-DISC)/2.
   20 IF(DABS(XAI)-ZER)31,31,32
   31 NR=NR+1
      X(NR)=XAR
      NR=NR+1
      X(NR)=XBR
   32 IF(N-3)8,8,12
   12 N=N-2
      IF(N-3)17,43,14
   14 DO 15 I=1,N
   15 A(I)=B(I+2)
      GO TO 7
   17 NR=NR+1
   41 X(NR)=-B(4)/B(3)
      GO TO 8
   43 P=-B(4)/B(3)
      Q=-B(5)/B(3)
      GO TO 9
    8 CONTINUE
```

```
FORTRAN IV G LEVEL  21          POLYS          DATE = 75205          12/48/58          PAGE 0002

 0C59              RETURN
 0060              END
```

METHODS OF RESPONSE SURFACE ANALYSIS

by

JOYCE LYNN TAYLOR

B. S., Oklahoma State University, 1973

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1975

In many fields of research the experimenter is concerned with finding a best set of operating conditions for a system. In response surface methodology the responses of this system are defined by a second-order model of the form

$$\eta = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \sum_{m>j} \beta_{jm} x_j x_m + \sum_{j=1}^{k} \beta_{jj} x_j^2 \quad .$$

The first part of this paper has divided the analysis of the response surface into three parts and discusses the computational techniques needed for a general second-order model in k independent variables. The first section is the regression analysis which estimates the coefficients of the response model and locates the point where the best set of operating conditions occurs. This point is called the stationary point. Second is the canonical analysis where one determines the nature and orientation of the response surface. Finally a $(1 - \alpha)100\%$ confidence region is determined for the stationary point. This confidence region not only gives an idea of the location of the optimum set of operating conditions, but it may also provide information about the location of future experiment points.

In the second part of this paper, examples of these techniques are shown using a second-order model with two independent variables. Output from a computer program which performs the three parts of the response surface analysis and plots the confidence region for the stationary point is included. The program itself is displayed in the appendix of the paper.